Annotations and Functional Analyses of the Rice *WRKY* Gene Superfamily Reveal Positive and Negative Regulators of Abscisic Acid Signaling in Aleurone Cells[1][w]

Zhen Xie[2], Zhong-Lin Zhang[2,3], Xiaolu Zou, Jie Huang, Paul Ruas, Daniel Thompson, and Qingxi J. Shen*

Department of Biological Sciences, University of Nevada, Las Vegas, Nevada 89154

The WRKY proteins are a superfamily of regulators that control diverse developmental and physiological processes. This family was believed to be plant specific until the recent identification of *WRKY* genes in nonphotosynthetic eukaryotes. We have undertaken a comprehensive computational analysis of the rice (*Oryza sativa*) genomic sequences and predicted the structures of 81 *OsWRKY* genes, 48 of which are supported by full-length cDNA sequences. Eleven OsWRKY proteins contain two conserved WRKY domains, while the rest have only one. Phylogenetic analyses of the WRKY domain sequences provide support for the hypothesis that gene duplication of single- and two-domain WRKY genes, and loss of the WRKY domain, occurred in the evolutionary history of this gene family in rice. The phylogeny deduced from the WRKY domain peptide sequences is further supported by the position and phase of the intron in the regions encoding the WRKY domains. Analyses for chromosomal distributions reveal that 26% of the predicted *OsWRKY* genes are located on chromosome 1. Among the dozen genes tested, *OsWRKY24*, *-51*, *-71*, and *-72* are induced by abscisic acid (ABA) in aleurone cells. Using a transient expression system, we have demonstrated that OsWRKY24 and -45 repress ABA induction of the *HVA22* promoter-β-glucuronidase construct, while OsWRKY72 and -77 synergistically interact with ABA to activate this reporter construct. This study provides a solid base for functional genomics studies of this important superfamily of regulatory genes in monocotyledonous plants and reveals a novel function for *WRKY* genes, i.e. mediating plant responses to ABA.

The *WRKY* genes encode a large group of transcription factors. There are over 70 *WRKY* genes in Arabidopsis (*Arabidopsis thaliana*; Eulgem et al., 2000; Dong et al., 2003) and rice (*Oryza sativa*; Goff et al., 2002; Zhang et al., 2004). This family is defined by a domain of 60 amino acids, which contains the amino acid sequence WRKY at its amino-terminal end and a putative zinc finger motif at its carboxy-terminal end. Some of the WRKY proteins contain two WRKY domains, while others have only one. Most of the published WRKY proteins bind to the cognate cis-acting element (C/T)TGAC(T/C) in the promoter (for review, see Eulgem et al., 2000) or the 5′ untranslated regions of target genes (Yu et al., 2001). For the WRKY proteins containing two WRKY domains, such as PcWRKY1 (Eulgem et al., 1999), SPF1 (Ishiguro and Nakamura, 1994), and AtZAP1 (de Pater et al., 1996), the C-terminal domain has the major DNA-binding activities.

WRKY proteins function as transcriptional activators and repressors. For example, the PcWRKY1 protein functions as a transcription activator binding to the W-box in the *PcWRKY1* promoter as shown in the assays with parsley (*Petroselinum crispum*) protoplasts and yeast (*Saccharomyces cerevisiae*) as LexA-WRKY1 fusion (Eulgem et al., 1999). Similarly, Arabidopsis ZAP1 binds to and activates a synthetic promoter containing the W-box in yeast and *Catharanthus roseus* suspension cells (de Pater et al., 1996). AtWRKY22 and AtWRKY29 transactivate the promoters of *WRKY29* and *FRK1* (*FLG22-INDUCED RECEPTOR-LIKE KINASE 1*), respectively, in the Arabidopsis protoplasts (Asai et al., 2002). Overexpression of *AtWRKY18* in transgenic Arabidopsis plants results in enhanced expression of pathogenesis-related genes (Chen and Chen, 2002). Recently, we have demonstrated that rice *WRKY71*, a homolog of wild oat (*Avena sativa* subsp. *fatua*) *ABF2* (Rushton et al., 1995), encodes a transcriptional repressor of GA signaling in aleurone cells (Zhang et al., 2004). Some *WRKY* genes can function as either a repressor or an activator. For instance, *AtWRKY6* suppresses its own promoter as well as the

[1] This work was supported by the U.S. Department of Agriculture (grant no. 02–35301–12066), by the National Institutes of Health (Biomedical Research Infrastructure Network seed grant no. P20 RR16464), and by the University of Las Vegas, Nevada (start-up funds to Q.J.S.). X.Z. was supported by a National Science Foundation Experimental Program to Stimulate Competitive Research (EPSCoR), Integrative Approaches to Abiotic Stress (EPS–0132556), graduate assistantship and Z.-L.Z. by a National Science Foundation EPSCoR Advanced Computing in Environmental Sciences postdoctoral fellowship.

[2] These authors contributed equally to the paper.

[3] Present address: Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720 and Plant Gene Expression Center, U.S. Department of Agriculture, Albany, CA 94710.

* Corresponding author; e-mail jeffery.shen@ccmail.nevada.edu; fax 702–895–3956.

[w] The online version of this article contains Web-only data.

Article, publication date, and citation information can be found at www.plantphysiol.org/cgi/doi/10.1104/pp.104.054312.

promoter of a closely related WRKY family member, whereas it activates the promoters of a receptor-like protein kinase (*SIRK*) and the senescence- and pathogen defense-associated *PR1* genes (Robatzek and Somssich, 2002).

WRKY genes play a variety of developmental and physiological roles in plants. The most reported studies for this superfamily of genes address their involvement in salicylic acid (SA) and disease responses (Chen and Chen, 2000; Dellagi et al., 2000; Du and Chen, 2000; Eulgem et al., 2000; Kim et al., 2000; Asai et al., 2002). In addition, *WRKY* genes are involved in plant responses to freezing (Huang and Duman, 2002), wounding (Hara et al., 2000), oxidative stress (Rizhsky et al., 2004), drought, salinity, cold, and heat (Pnueli et al., 2002; Rizhsky et al., 2002; Seki et al., 2002). Some *WRKY* genes regulate embryogenesis (Lagace and Matton, 2004), seed coat and trichome development (Johnson et al., 2002), and senescence (Hinderhofer and Zentgraf, 2001; Robatzek and Somssich, 2001). Biosynthesis of anthocyanin (Johnson et al., 2002), starch (Sun et al., 2003), and sesquiterpene (Xu et al., 2004) are also dependent on WRKY proteins. *WRKY* genes may control seed germination and postgermination growth as well, because two wild oat WRKY proteins (ABF1 and ABF2) bind to the box2/W-box of the GA-regulated *α-Amy2* promoter (Rushton et al., 1995) and *OsWRKY71* encodes a transcriptional repressor of GA signaling in aleurone cells (Zhang et al., 2004). The report of *AtWRKY70* in mediating both SA and jasmonic acid responses indicates the importance of this gene in the cross-talk of hormone signaling (Li et al., 2004).

Previously, we reported the annotations of 77 *OsWRKY* genes (Zhang et al., 2004). Here we present more *OsWRKY* genes, along with the bioinformatic analysis data such as their chromosomal localizations, intron-exon structures, full-length cDNA (FL-cDNA) support, alternative splicing, phylogenetic relationships, and the related protein motifs for these 81 *OsWRKY* genes. Our studies also reveal that there are at least four *OsWRKY* genes involved in abscisic acid (ABA) signaling in aleurone cells, with two functioning as positive and two as negative regulators. To our knowledge, this is the first report of such activity by a WRKY protein.

RESULTS

Identification of 81 *WRKY* Genes in the Rice Genome

A systematic analysis was carried out to identify *WRKY* genes in the rice genome using the publicly available genomic sequences of rice L. subsp. *japonica* (Goff et al., 2002) and subsp. *indica* (Yu et al., 2002) that were downloaded from GenBank and from the Web site of the Beijing Genomics Institute (http://btn.genomics.org.cn/rice). Using GENSCAN, a gene prediction program (Burge and Karlin, 1997), we constructed databases of the predicted rice coding regions and peptides (Zhang et al., 2004). *WRKY* genes were identified by using the HMMsearch program and a Hidden Markov Model (PF03106; http://pfam.wustl.edu), with a cutoff E-value of 0.1. A total of 158 *WRKY* genes were identified in *japonica* and *indica* rice genome sequences. Manual inspections eliminated 32 candidate genes because the deduced proteins for these genes do not contain any of the conserved amino acid residues, W(R/K)(K/R)Y, in the N terminus of the WRKY domain. The corresponding coding region sequences (CRSs) of the remaining 126 sequences were analyzed for redundant sequences by performing pairwise comparisons. An arbitrary cutoff value of 95% was used to take into account sequencing errors and natural variations in the *japonica* and *indica* rice genomes. Thus, if the identity between 2 CRSs (including both exon and intron sequences between the start and stop codons) is equal to or higher than 95%, they are considered to be redundant and only 1 is kept. This analysis eliminated 54 sequences. The WRKY domain peptide sequences of the remaining 72 putative *WRKY* genes were aligned using ClustalW (Thompson et al., 1994). The alignment output was used to build a rice-specific WRKY HMM model using HMMbuild and HMMcalibrate in the HMMER package. Scanning with this new model against the rice peptide databases (Zhang et al., 2004), also with a cutoff E-value of 0.1, resulted in 178 candidates. Twenty-nine of these candidate sequences were eliminated because they did not contain the conserved WRKY motif. One gene identified in Contig96377 (accession no. AAAA01094809.1) was eliminated because it was very short, missing both the 5′ and 3′ regions of the gene. The CRSs of the remaining 148 candidates were retrieved and a redundancy analysis was carried out as described, eliminating 67 sequences (see Supplemental Table I). In total, 81 *WRKY* genes were identified, with 45 in the *japonica* genome sequences and 36 in the *indica* genome sequences. The detailed information on these 81 *OsWRKY* genes is listed in Supplemental Table II.

Manual Inspections Reveal Annotation Errors and Alternative Spliced Open Reading Frames

Approximately 36% of Arabidopsis genes identified through the automated predictions contained errors (Haas et al., 2002; Meyers et al., 2003). We took several approaches to refine the accuracy of the collected *OsWRKY* genes. First of all, three genes (*OsWRKY41*, *-43*, and *-44*) were reannotated using FGENESH (www.softberry.com), because the first introns of these genes were too small (Supplemental Table III). Second, two partial genes published previously (Zhang et al., 2004), *OsWRKY46* (missing both the 5′ and 3′ ends) and *OsWRKY63* (missing the 3′ end), were replaced with full-length genes we identified in newly available bacterial artificial chromosome (BAC) sequences. However, no new sequences for *OsWRKY50* have been identified yet and hence it is still short at the 3′ end. Third, the predicted genes were analyzed by

comparison to DNA and deduced amino acid sequences of the rice FL-cDNA sequences (Kikuchi et al., 2003) using BLAST analyses. An annotation is considered to have a significant FL-cDNA match if (1) the predicted coding sequence is ≥95% identical to the corresponding region of its FL-cDNA sequence at both the DNA and deduced amino acid sequence levels; and (2) the length of the identical region between the predicted and FL-cDNA sequences is ≥95% of the length of the entire sequence at both the DNA and deduced amino acid sequence levels. In contrast, a cDNA is considered as a partial match if the identity is ≥95%, but the length of the identical region is between 50% and 94% (Supplemental Table IV). Among the 81 predicted *OsWRKY* genes, 36 had significant matches. This analysis also revealed that 17 predicted sequences had annotation errors, including incorrect intron-exon splice boundaries, incorrect number of exons, and wrong start/stop codons (Supplemental Table III). After manual corrections, 48 *OsWRKY* genes had significant matches, two had partial matches, and 31 did not have any matches. The accession numbers of the matched FL-cDNA sequences are listed in Supplemental Table II.

Recently, it has been reported that about 27% of 18,933 transcription units in rice genomes contain 2 or more transcripts (Kikuchi et al., 2003) due to alternative splicing, which results in alternative open reading frames (ORFs) that differ in initiation and termination sites, as well as splice donor and acceptor sites. To date, three Arabidopsis *WRKY* genes (*AtWRKY1*, *-4*, and *-43*) have been identified as having two ORFs. Because 66 out of 81 *OsWRKY* genes in our collection have more than 2 exons, we tested whether they contain alternative ORFs. In this study, we found that five *OsWRKY* loci (*OsWRKY1*, *-8*, *-35*, *-39*, and *-57*) contain alternative structures (Fig. 1). In two of these *OsWRKY* genes, *OsWRKY1* and *OsWRKY57*, one of the alternative peptides does not contain a WRKY domain. In another one (*OsWRKY35*), the alternative splice variant retains part (54 out of 88 nucleotides) of the fourth intron located in the C-terminal WRKY domain, resulting in the addition of 18 amino acids to the deduced peptide. The predicted gene structure of *OsWRKY39* is supported by a FL-cDNA (accession AK066775). In addition, an alternative structure (variant 1) is revealed by another FL-cDNA (AK119593). This splice variant retains the first intron of the predicted structure, which introduced a stop codon right after the 127th codon, thereby eliminating the WRKY domain that lies downstream of codon 127. *OsWRKY34* has a significant FL-cDNA match (AK072906), but it retains the third predicted intron. No FL-cDNA that completely matches the predicted structure has been found yet. These results suggest that alternative splicing might be involved in modulating the function of WRKY proteins.

OsWRKY Proteins Are Classified into Four Subgroups Based on Their WRKY Domain Sequences

The most prominent structural feature of these proteins is the WRKY domain of 60 amino acid residues. Among 93 WRKY domains found in WRKY

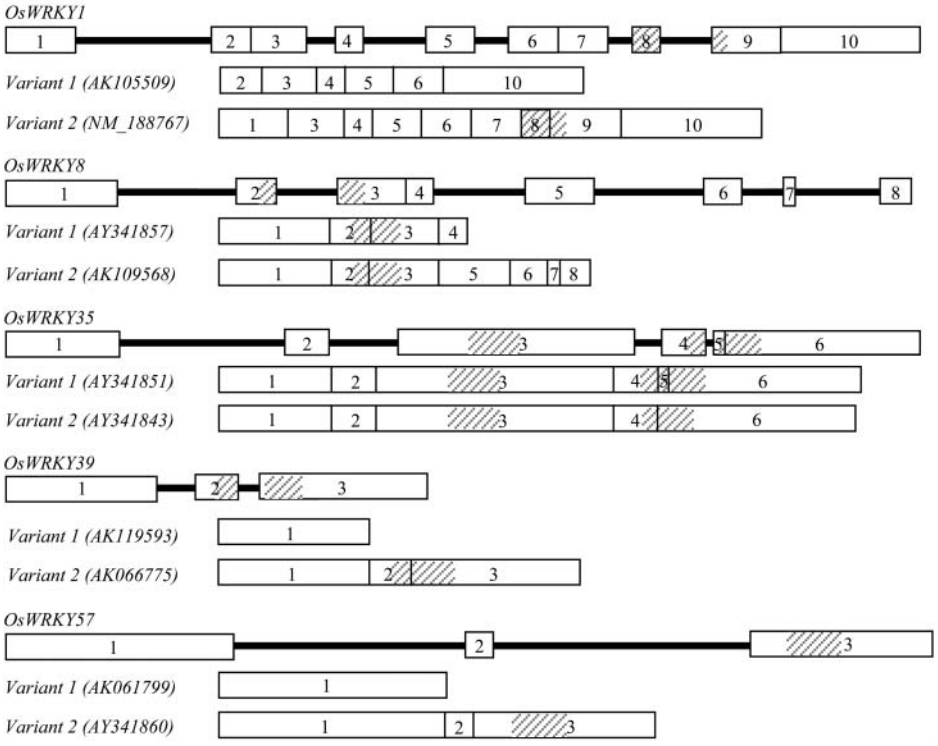

Figure 1. Schematic diagrams of alternative spliced gene structures. Accession numbers of the significantly matched FL-cDNA sequences are shown in parentheses. Rectangles with numbers represent exons; thick lines represent introns; hatched regions represent WRKY domains. The diagrams were drawn to scale.

proteins, 87 domains contain the highly conserved WRKY motif, including 77 with WRKYGQK, 6 with WRKYGKK, and 4 with WRKYGEK. However, 6 domains in OsWRKY41, -60, -61, -63, and -81 have either 1 or 2 mismatched amino acids within the WRKY motif (referred to as an atypical WRKY motif). Thus, we suggest that W(R/K)(K/R)Y be considered as a new consensus WRKY motif. Similar to the Arabidopsis WRKY proteins (Eulgem et al., 2000), OsWRKY proteins were classified into 4 groups (Fig. 2): 11 members with 2 WRKY domains belong to group I; the WRKY domains of 46 group II members have a C_2H_2-type (C-$X_{4–5}$-C-$X_{22–24}$-H-$X_{1–2}$-H) zinc finger motif; 19 WRKY proteins belong to group III with a C_2HC-type (C-$X_{6–7}$-C-$X_{23–33}$-H-X_1-C) zinc finger motif; and 5 group IV WRKY proteins contain no zinc finger motif. The zinc finger motif patterns in group II and III OsWRKY proteins are slightly different from those in group II (C-$X_{4–5}$-C-$X_{22–23}$-H-X_1-H) and group III (C-X_7-C-X_{23}-H-X_1-C) of Arabidopsis WRKY proteins, respectively (Eulgem et al., 2000). In addition, based on their zinc finger motifs, members in groups I and IV can be further divided into four subgroups (Ia, Ib, IVa, and IVb). Group Ia proteins contain a C_2H_2-type zinc finger motif with a common WRKY motif; group Ib proteins contain a C_2HC-type zinc finger motif, usually with an atypical WRKY motif; group IVa proteins have only a partial zinc finger motif (CX_4C); and group IVb proteins do not contain the conserved Cys or His residues at all. With the exception of *OsWRKY56*, which is supported by a FL-cDNA sequence, group IV genes might contain annotation errors caused by genomic sequencing or gene prediction software, or they might be pseudogenes with no biological functions.

Phylogenetic Analysis of the *OsWRKY* Gene Family

To study the phylogenic relationship among the *OsWRKY* genes and generate an evolutionary framework of this gene family, a multiple peptide sequence alignment of all 93 WRKY domains, including both C-terminal and N-terminal domains of group I genes, was constructed using ClustalW (Thompson et al., 1994). For the *WRKY* genes that have alternatively spliced ORFs, only variant 2 ORFs of these genes (*OsWRKY1v2*, *-8v2*, *-35v2*, and *-57v2*) were included in the phylogenetic analysis, because the sequences and structures of variant 2 ORFs were more conserved than those of variant 1 ORFs. The sequence alignment output was used to reconstruct the unrooted majority rule consensus phylogenetic tree within the Bayesian framework (Fig. 3). Although the basal nodes of the tree are not well resolved, there are several patterns visible in the branches with 95% to 100% Bayesian posterior probabilities, as described in the next paragraph.

Recently, it was reported that group I *WRKY* genes may represent the ancestral form of *WRKY* genes (Ulker and Somssich, 2004). The loss of one WRKY domain in the origin of several group II genes (Fig. 3, names in black) that evolved from group Ia genes (Fig. 3, names in red) is evident in the separate monophyletic clusters of all C- and N-terminal WRKY domains. The C-terminal WRKY domain branch includes genes that have lost the N-terminal domain, such as two group IVa genes (*OsWRKY33* and *-38*) and a group II gene (*OsWRKY80*). In contrast, in the C-terminal WRKY domain cluster, a group II gene (*OsWRKY57v2*) appears to be evolved from group Ia genes by loss of the C-terminal domain. Of the remaining 44 group II genes, 40 cluster in 5 well-supported branches that share an unresolved basal node with the group Ia genes, suggesting that either group II genes evolved from the ancestral form of group Ia genes or two-domain group Ia genes evolved from gene duplication of group II genes. Both WRKY domains of the group Ib genes (Fig. 3, names in purple) are imbedded within a cluster of group III genes (Fig. 3, names in green). Although these relationships are not as well supported as the patterns described for group Ia, the N-terminal WRKY domains of group Ib genes form a monophyletic cluster and the C-terminal WRKY domains of group Ib genes connect basally to form a monophyletic cluster of all Ib genes and 4 group III genes with 82% Bayesian posterior probability, suggesting a WRKY domain duplication event of a group III gene gave rise to group Ib genes. The three group IVa *WRKY* genes (*OsWRKY52*, *-56*, and *-58*; Fig. 3, names in blue) are scattered throughout the tree and branch, with low resolution, from group II and III genes.

Intron Phases and Positions Support the Phylogeny of *OsWRKY* Genes

The position and phase of the intron in each region encoding the WRKY domain provides additional support to the phylogeny in Figure 3. An intron position refers to a codon into which an intron is inserted, whereas an intron phase refers to where in a codon an intron is inserted: immediately upstream (phase 0), after the first nucleotide (phase 1), or after the second nucleotide (phase 2) of a codon. About 71% (66 out of 93) of WRKY domains found in *OsWRKY* genes contain an intron in a conserved position, including both WRKY domains of group Ib genes (Fig. 2, red boxes and Fig. 3, right). This phase 2 intron is localized at the 11th codon downstream of the WRKY motif, interrupting the codon encoding Arg (91% of 66 genes), Ser (5%), Lys (3%), or Asn (1%; Fig. 2). The presence of the intron in both WRKY domains of group Ib genes and in all other gene groups suggests that its origin precedes much of the gene duplication in this family. Of the 20 WRKY domains that do not have this conserved intron, 8 are in the monophyletic cluster of N-terminal WRKY domains of group Ia genes, 3 domains are in a monophyletic cluster of group II genes (*OsWRKY1v2* to *OsWRKY 62*; Fig. 3), and 5 domains share a branch tip among group II genes (*OsWRKY2, -10, -25, -31*, and *-44*; Fig. 3). The

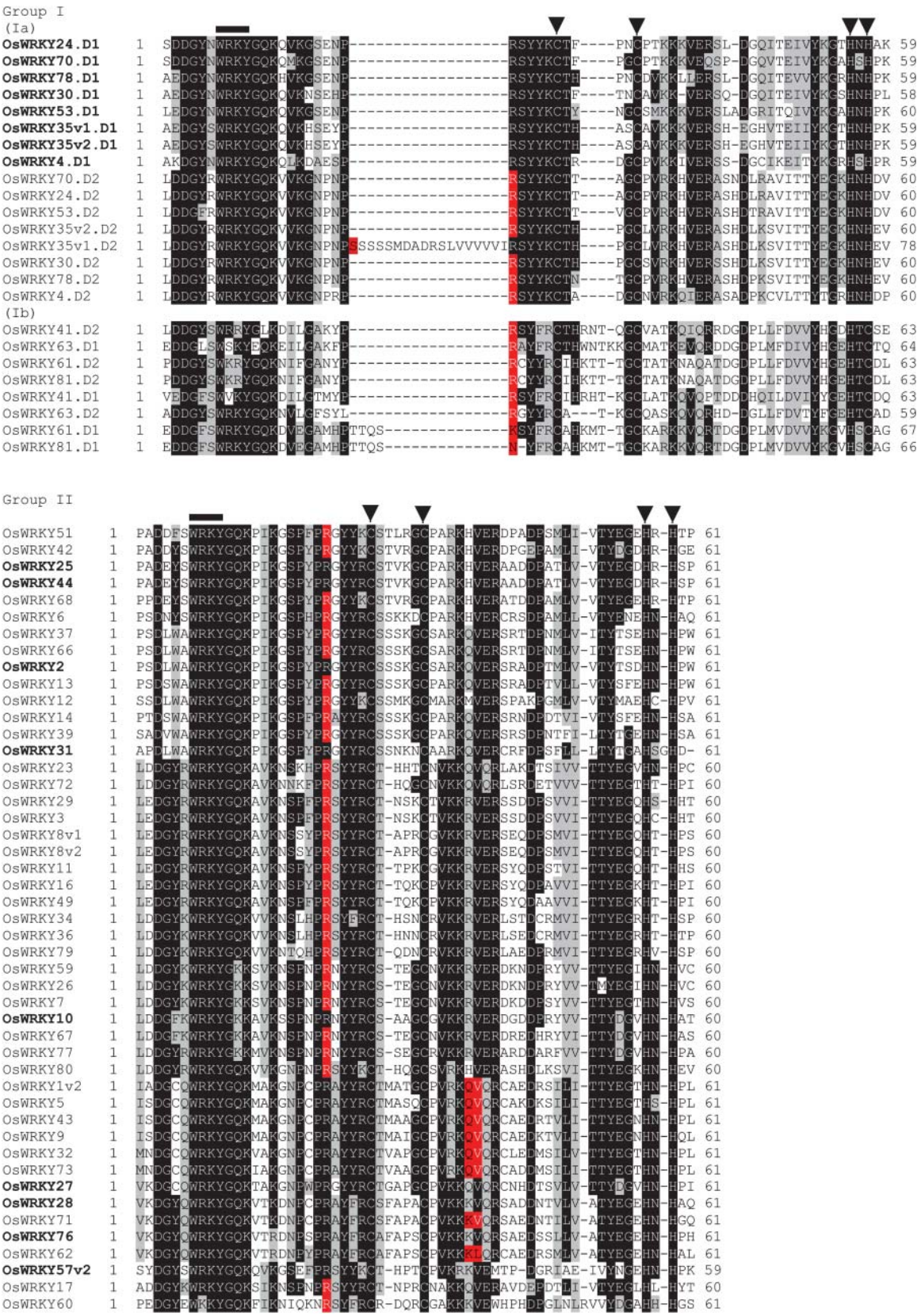

Figure 2. Multiple sequence alignment of WRKY domains. Eighty-one OsWRKY proteins were divided into four groups based on the features of their WRKY domains. Groups I and IV were further divided into subgroups (Ia, Ib, IVa, and IVb). The names in bold highlight the WRKY proteins that do not contain an intron within the region coding for the WRKY domains. WRKY motifs are below the black line; amino acid residues potentially interacting with zinc ligands are indicated with arrows. Letters in red represent the amino acid residues whose codons are interrupted by introns.

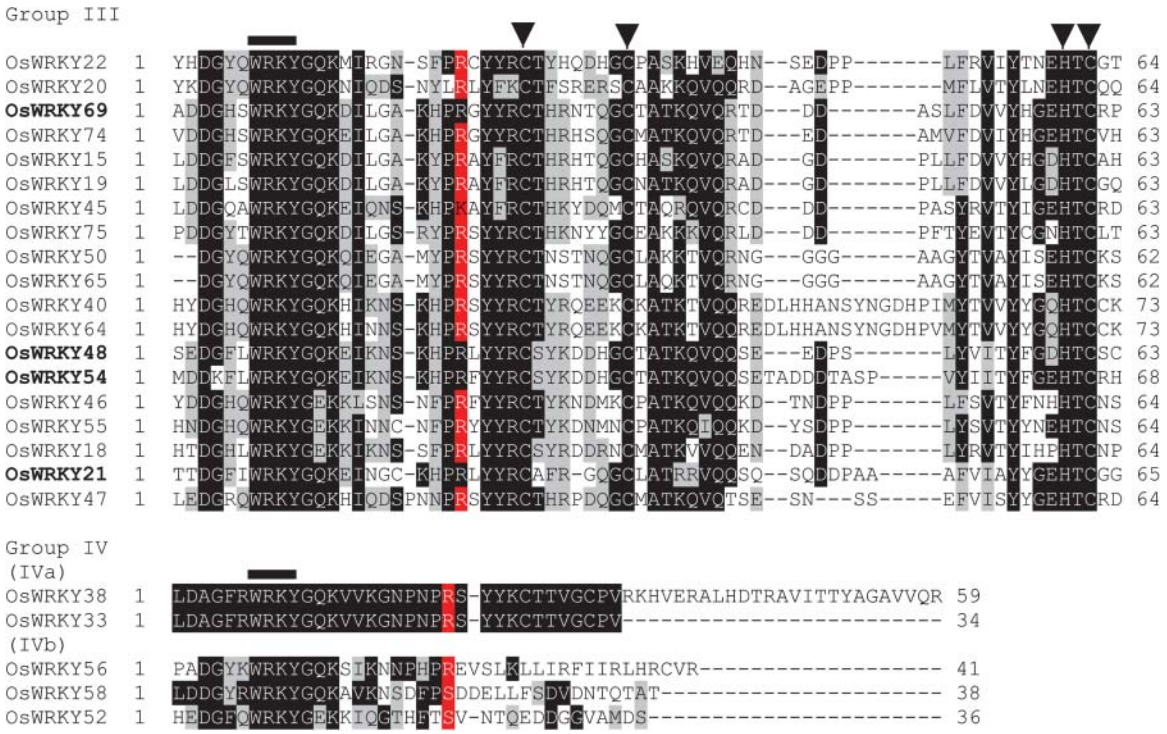

Figure 2. (*Continued.*)

remaining domains that lack the conserved intron belong to group III genes (*OsWRKY21*, *-48*, *-54*, and *-69*; Fig. 3). If this intron is ancestral to most gene duplication events in the *WRKY* gene family, only seven independent losses are required to account for the intron distribution. The absence of this phase 2 intron from the N-terminal WRKY domains of all group Ia proteins is another feature that distinguishes group Ia proteins from group Ib proteins (Fig. 2). Another type of intron within the WRKY domain occurs only in a group II gene cluster. This phase 0 intron is localized between the 28th and 29th codons downstream of the WRKY motif, and is found only in eight group II *WRKY* genes, including *OsWRKY1v2*, *-5*, *-9*, *-32*, *-43*, *-62*, *-71*, and *-73*. The 28th and 29th codons were conserved, encoding Gln or Lys and Val or Leu, respectively (Fig. 2).

Chromosomal Localization of *OsWRKY* Genes

Forty-five *WRKY* genes revealed in the BAC sequences of the *japonica* genome were localized on chromosomes based on the BAC information obtained from National Center for Biotechnology Information (NCBI). To investigate the chromosomal localizations of *OsWRKY* genes identified in the *indica* genome, in which draft sequence was produced by whole-genome shotgun sequencing (Yu et al., 2002), *WRKY* genes containing BAC sequences were used for BLAST analyses against the *japonica* genome sequence. Twenty *OsWRKY* genes found in the *indica* genome were also identified in the *japonica* genome and hence localized on chromosomes. The detailed chromosomal information of *OsWRKY* genes is listed in Supplemental Table V. *OsWRKY* genes in this collection were distributed across all 12 chromosomes (Fig. 4A). Interestingly, 21 of the 81 *OsWRKY* genes were found on chromosome 1, including 2 tandem repeats of *OsWRKY15*. In other words, approximately 26% of the identified *OsWRKY* genes are on chromosome 1, yet the length of chromosome 1 only accounts for 11% of the rice genome. These results suggest that there might be *WRKY* gene hot spots in the rice genome, with the highest density on chromosome 1. Based on the sequence of chromosome 1, the physical positions of 21 *OsWRKY* genes were determined by BLAST analyses and are shown in Figure 4B. Eighteen of the 21 *OsWRKY* genes were clustered into 6 regions (A–F) on chromosome 1. Most of the *OsWRKY* genes within the same region are highly conserved. Furthermore, three possible tandem gene duplication events were revealed in regions D, E, and F because those genes were clustered on the deepest branches of the tree, suggesting a close phylogenetic relationship (Fig. 3). It will be interesting to analyze whether these regions with a higher density of *OsWRKY* genes are in genome tandem arrays or intrachromosomally and interchromosomally duplicated areas.

Motif Analyses of OsWRKY Proteins

Identification of the conserved protein motifs might help elucidate protein functions. It has been suggested that several Arabidopsis group IIb WRKY proteins contain a conserved Leu zipper motif (Eulgem et al.,

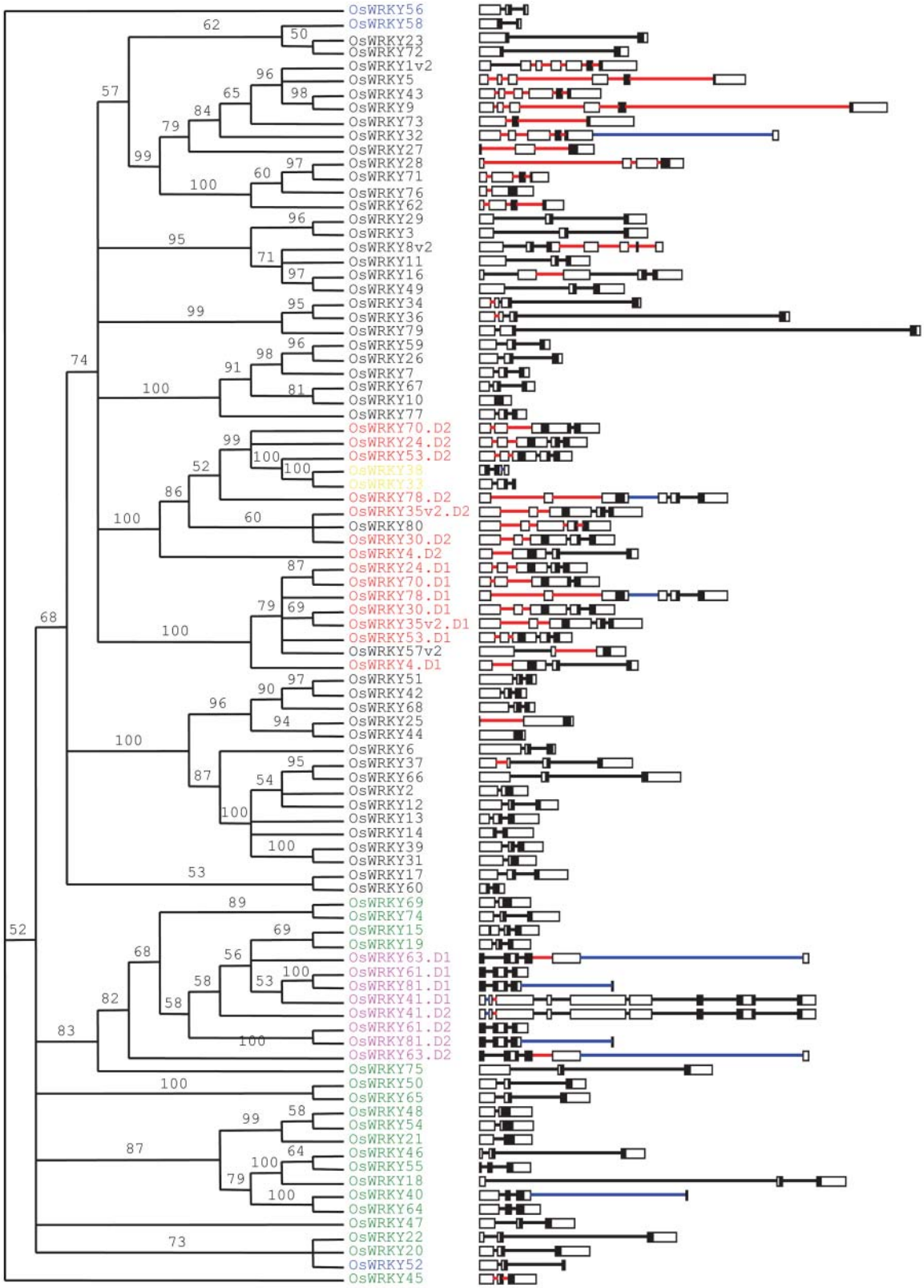

Figure 3. Phylogenetic analyses of OsWRKY proteins. The majority rule consensus phylogenetic unrooted tree was reconstructed based on the WRKY domain peptide sequences by the MrBayes 3.0 program (Huelsenbeck et al., 2000). Statistical support values are shown above each node. D1 represents the N-terminal WRKY domain of group I proteins; D2 represents the C-terminal WRKY domain of group I proteins. Group Ia proteins (red); group Ib proteins (purple); group II proteins (black), group III proteins (green); group IVa proteins (yellow); group IVb proteins (blue). The intron patterns for 81 *OsWRKY* genes were drawn to scale and are shown beside the phylogenetic tree. White rectangles represent exons; black rectangles denote WRKY domain regions. Lines drawn in red represent phase 0 introns; lines in blue represent phase 1 introns; lines in black represent phase 2 introns.

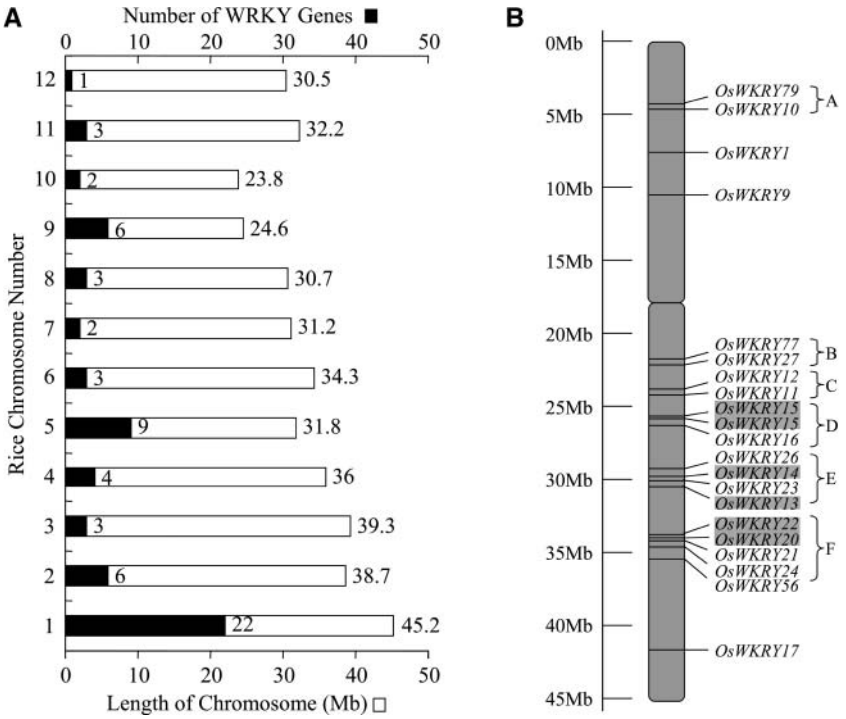

Figure 4. Chromosomal distribution of *OsWRKY* Genes. A, The distribution of *OsWRKY* genes in the rice genome. The chromosomal information for *OsWRKY* genes identified in the *japonica* genome was determined based on BAC information obtained from the NCBI (http://www.ncbi.nlm.nih.gov). To determine chromosomal information for genes identified in the *indica* genome, the corresponding contig sequences were used for BLAST search against the *japonica* BAC sequences in the NCBI. The overlapped sequences were determined based on BLAST results, and the chromosomal information on these genes was obtained from the overlapped *japonica* BAC information. Detailed results were listed in Supplemental Table V. The information on the length of rice chromosomes was obtained from the IRGSP (http://rgp.dna.affrc.go.jp/IRGSP/). B, Chromosomal locations of *OsWRKY* genes in rice chromosome 1. Six regions (A–F) with a higher density of *OsWRKY* genes were revealed. Possible duplicated *OsWRKY* genes are highlighted in gray.

2000), which is characterized by a region containing several repeats of seven hydrophobic residues with a strong preference for Leu at every seventh position, although other small hydrophobic amino acids are also acceptable (Pu and Struhl, 1991; Acharya et al., 2002). Among 81 OsWRKY proteins, 4 group II proteins (OsWRKY1, -9, -43, and -71) contain Leu zipper motifs (Supplemental Fig. 1). Experimental evidence is necessary to determine whether these proteins can indeed dimerize. Another notable motif, the HARF motif, is found in group IId AtWRKY proteins, although its function is still unknown (Eulgem et al., 2000). This HARF motif was only found in a few group II OsWRKY proteins (OsWRKY25, -44, -51, and -68; Fig. 3). OsWRKY27 and -80 contain a hydroxylation motif, which is present in proteins such as HIF-1α (hypoxia-inducible factor 1α) and is involved in sensing oxygen levels in cells (Huang et al., 2002). Recently, sumoylation is shown to be involved in both positive and negative regulations of transcription by modulating the ability of transcription factors to interact with their partners, alter their patterns of subcellular localization, and control their stability, leading to positive or negative regulation of transcription (for review see Seeler and Dejean, 2003; Verger et al., 2003). The consensus sequence of the sumoylation motif is ΨKXD/E, where Ψ is often an aliphatic residue but sometimes can be basic or aromatic amino acids such as Lys and Phe, respectively (Zhou et al., 2004). We found that 35 OsWRKY peptide sequences contained one or more sumoylation motif(s). Furthermore, four WRKY proteins (OsWRKY1, -8, -35, and -57), which have alternative ORFs, contain one or more sumoylation motifs. It will be interesting to determine whether these proteins can be sumoylated and whether the function of these proteins can be altered through this modification.

WRKY proteins can function as activators and repressors of many different biological processes (de Pater et al., 1996; Eulgem et al., 1999; Asai et al., 2002; Chen and Chen, 2002). Some can function as activators in one pathway but as repressors in another (Robatzek and Somssich, 2002). To predict putative activating and repressing activities of OsWRKY proteins, the deduced peptide sequences were searched for the consensus coactivator motif, LXXLL, where L is Leu and X is any amino acid (Savkur and Burris, 2004) and the active repressor motif, LXLXLX (Tiwari et al., 2004). Sixteen OsWRKY proteins (OsWRKY2, -11, -17, -20, -21, -36, -40, -41, -45, -48, -50, -54, -57, -65, -73, and -75) contain the coactivator motif, and nine proteins (OsWRKY27, -28, -29, -32, -40, -51, -64, -71, and 73) contain the active repressor motifs. Moreover, OsWRKY40 and -73 contain both coactivator and active repressor motifs. The motif structures of OsWRKY24, -45, -51, -71, -72, and -77, whose functions were experimentally analyzed in aleurone cells as detailed below, are shown in Figure 5.

Expression Patterns of *OsWRKY* Genes in Rice Aleurone Cells

To investigate *WRKY* genes expressed in rice aleurone cells, northern-blot analyses were carried out using total RNA isolated from aleurone cells derived from rice embryoless half-seeds treated with or without the hormones, GA_3, ABA, or both. Steady-state

mRNA levels for several *OsWRKY* genes were determined and compared to those of *RAmy1A* (a GA-inducible α-amylase gene; Karrer et al., 1991), *RAB21* (an ABA-inducible gene; Mundy and Chua, 1988), and *OsActin1* (a constitutively expressed rice gene; McElroy et al., 1990). As shown in Figure 6, the expression of *RAmy1A* and *RAB21* was strongly induced after GA_3 and ABA treatments at 24 and 48 h, respectively. The steady-state mRNA levels of *OsActin1* were similar in all samples, suggesting that an equal amount of RNA was used on the blots. The steady-state mRNA signals of 5 *WRKY* genes, *OsWRKY24*, *-45*, *-51*, *-71*, and *-72*, were detectable in rice aleurone cells. *OsWRKY24*, *-51*, and *-71* transcripts were relatively abundant compared to those of *OsWRKY45* and *-72*. The low level of *OsWRKY45* expression was not affected by either ABA or GA_3, but OsWRKY72 demonstrates a clear induction upon treatment with ABA (Fig. 6). The abundant transcripts of *OsWRKY24*, *-51*, and *-71* were enhanced upon ABA treatment. Furthermore, the mRNA levels of *OsWRKY51* and *-71* were down-regulated in response to exogenous GA_3 treatments at 24- and 48-h time points. *OsWRKY24* was slightly down-regulated by GA at 4- and 24-h time points.

We also examined OsWRKY77 by northern-blot analysis but could not detect a signal. The signal of *OsWRKY77* transcription, however, can be detected by reverse transcription-PCR (Zhang et al., 2004).

Functional Analyses of *OsWRKY* Genes in Aleurone Cells

Because at least four out of six *OsWRKY* genes (*OsWRKY24*, -45, *-51*, *-71*, *-72*, and -77) expressed in aleurone cells are induced by ABA, we investigated whether they are involved in ABA signaling by transient expression of these *OsWRKY* genes in aleurone cells. The genomic sequences of these six *OsWRKY* genes were cloned into an expression vector driven by the maize (*Zea mays*) ubiquitin promoter (Fig. 7A). We have previously demonstrated that the genomic coding sequences (including all exons and introns between the start and stop codons) of several effector genes, including *PKABA1* and *GAMYB*, function as well as their corresponding cDNA clones in the transient expression system (Zhang et al., 2004). The ABA-inducible reporter construct, *HVA22-GUS* (Shen and Ho, 1995), was used to study the function of OsWRKY proteins in ABA signaling. The *OsWRKY* effector constructs, along with the *HVA22-GUS* reporter and *UBI-LUC* internal control constructs, were introduced into barley (*Hordeum vulgare*) aleurone cells by particle bombardment. The *HVA22-GUS* reporter is induced 31-fold in the presence of ABA (Fig. 7B). Transient expression of *UBI-OsWRKY71* had little effect on ABA induction of *HVA22-GUS* (Zhang et al., 2004). However, OsWRKY24 and -45 repressed ABA induction of *HVA22-GUS*. OsWRKY24 almost completely blocked ABA induction of *HVA22-GUS*, while transient expression of *UBI-OsWRKY45* partially inhibited ABA-dependent induction of *HVA22-GUS*. In contrast, coexpression of *UBI-OsWRKY72* and *UBI-OsWRKY77* induced the expression of *HVA22* to 4- and 11-fold, respectively, in the absence of ABA. With the ABA treatment, the induction levels were increased to 58- and 53-fold, respectively. These data suggest that OsWRKY72 and *-77* synergistically interact with ABA in activating the *HVA22* promoter in aleurone cells. Finally, transient expression of *UBI-OsWRKY51* had little effect on ABA induction of *HVA22-GUS*.

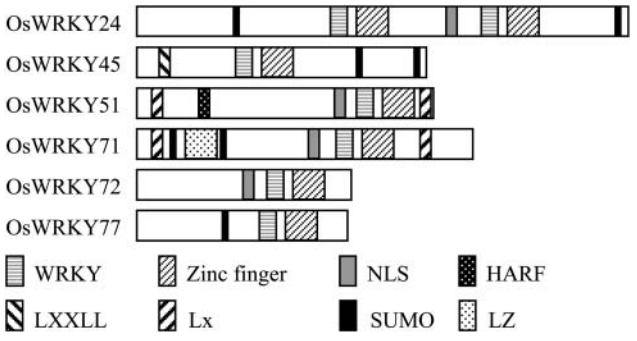

Figure 5. Schematic diagrams of conserved motifs found in the OsWRKY peptide sequences. Conserved motifs of OsWRKY24, -45, -51, -71, -72, and -77 were listed above. Other OsWRKY proteins were shown in Supplemental Figure 1. WRKY, WRKY motif; zinc finger, zinc finger motif; NLS, putative nuclear localization motif; HARF, conserved motif only found in a subgroup of WRKY proteins; LXXLL, putative coactivator motif; Lx, putative active repressor motif with the LXLXLX consensus sequences; SUMO, putative sumoylation motif; LZ, putative Leu-zipper motif.

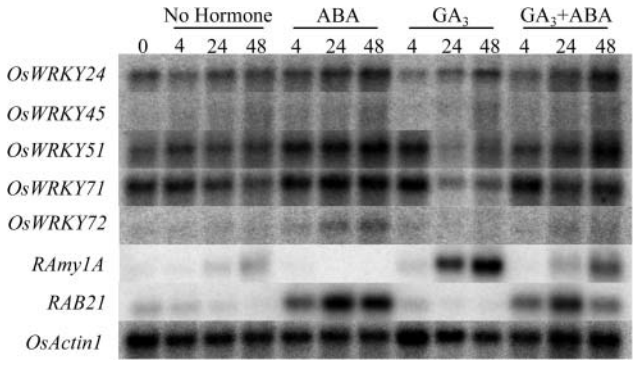

Figure 6. Expression patterns of *OsWRKY* genes in rice aleurone cells. After being imbibed for 2 d, about one-tenth of rice embryoless half-seeds were harvested as a control (0 h), and the rest were further treated without hormone (control) or with 1 μM GA_3, 20 μM ABA, or 1 μM GA_3 plus 20 μM ABA at 26°C for 4, 24, and 48 h. Total RNA was isolated from each treated rice embryoless half-seed. The northern blot, containing 10 μg of total RNA per sample, was hybridized with α-^{32}P-labeled gene-specific probes. To ensure equal loading of RNA, one representative blot was subsequently stripped and rehybridized with a probe against *OsActin1*.

DISCUSSION

This article reports the identification of 81 Os*WRKY* genes (Supplemental Table II) by reiterative

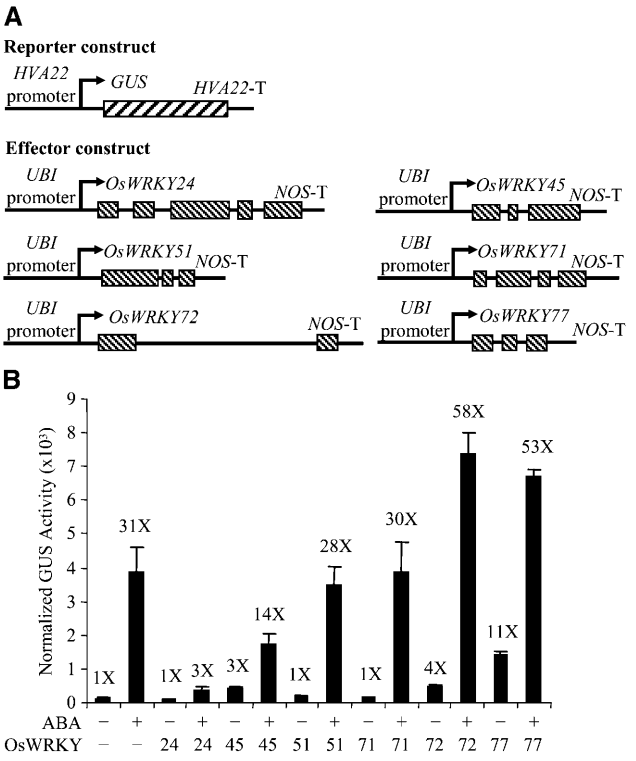

Figure 7. Functional analyses of OsWRKY proteins in ABA signaling. A, Schematic diagrams of the reporter and effector constructs used in the cobombardment experiments. The gene structures were drawn to scale. Shaded rectangles represent exons; lines between exons represent introns. B, Functions of OsWRKY proteins in ABA signaling. The reporter construct, *HVA22-GUS*, and the internal control construct, *UBI-LUC*, were cobombarded into barley half-seeds either with (+) or without (−) the effector construct by using an equal molar ratio of effector and reporter constructs. GUS activity was normalized in every independent transformation relative to the luciferase activity. Bars indicate GUS activities ±SE after 24 h of incubation of the bombarded half-seeds with (+) or without (−) 20 μM ABA. Data are means ± SE of four replicates.

bioinformatic analyses with the GENSCAN program followed by (1) removing redundant *OsWRKY* genes identified in the *japonica* and *indica* genome sequences (Supplemental Table I); (2) verification with FGENESH and manual inspections (Supplemental Table III); and (3) searching for supporting FL-cDNA sequences (Supplemental Tables II and IV). Most of the predicted *OsWRKY* genes are supported by FL-cDNA sequences, as demonstrated by the vigorous analyses that also led to the identification of *OsWRKY* mRNA variants resulting from alternative splicing (Fig. 1). Sequence alignment studies allowed the classification of OsWRKY proteins into four groups, which differ in the number of WRKY domains in the proteins and in the sequences of the WRKY and zinc finger motifs (Fig. 2). Phylogenetic analyses revealed both loss and gain of the WRKY domain in the evolutionary history of this gene family in rice (Fig. 3). Although *OsWRKY* genes have been found on each of the 12 rice chromosomes, they are not evenly distributed. Instead, chromosome 1 contains more than 25% of *OsWRKY* genes identified so far (Fig. 4). We demonstrated that *OsWRKY51* and *-71* were enhanced by ABA and suppressed by GA (Fig. 6). Finally, our studies identified two *OsWRKY* genes as negative regulators and two other *OsWRKY* genes as positive regulators of ABA signaling (Fig. 7).

The report demonstrates that the manual reannotation of genes is an essential component of genome analysis because of the limitations inherited in automated annotations and the complex nature of gene structures. This has also been demonstrated in the analysis of Arabidopsis *R* genes (Meyers et al., 2003). Our study shows that about 36% of predicted *OsWRKY* genes that have cDNA support contain errors of one type or another, assuming that all alternative forms of the cDNAs have been identified (Supplemental Table III). This assumption, of course, is likely to be an underestimation. The most frequent error we found was the wrong annotation of start and stop codons, which accounted for approximately 65% of the misannotations. The next frequent categories were the incorrect number of exons (approximately 18%) and incorrect placement of intron-exon splice boundaries (approximately 18%). FL-cDNA sequences are essential for the reannotation process; 64% of the predicted *OsWRKY* genes reported herein are supported by FL-cDNA sequences.

The *OsWRKY* genes encode proteins containing one or two WRKY domains. These two classes of *WRKY* genes may have arisen by gene duplication of the one-domain proteins, deletion of one WRKY domain in a two-domain protein, or a combination of both, i.e. deletion followed by duplication. Recently, *WRKY* expressed sequence tags have been identified from lower plants such as ferns (*Ceratopteris richardii*), mosses (*Physcomitrella patens*), and green algae (*Chlamydomonas reinhardtii*), and from the slime mold (*Dictyostelium discoideum*) and the unicellular protist (*Giardia lamblia*; Ulker and Somssich, 2004). Group I-like *WRKY* genes are present in the lower organisms, implying that the group I-like *WRKY* genes may represent the ancestral form of the WRKY family that originated approximately 1.5 to 2 billion years ago in eukaryotes, before the divergence of the plant phyla (Ulker and Somssich, 2004). However, rice group I *WRKY* genes can be further classified into two subgroups (group Ia and Ib). Two obvious features can distinguish group Ia from group Ib *WRKY* genes in rice. First, the zinc finger motif in both the N- and the C-terminal WRKY domain in group Ia genes is C_2H_2 type, while the zinc finger motif in both the N- and the C-terminal domain in group Ib genes is C_2HC type. Second, the C- but not the N-terminal domain of group Ia genes contains the conserved intron in the region encoding the WRKY domain. In contrast, both the N- and the C-terminal WRKY domains of the group Ib genes contain the conserved intron (Fig. 2). By comparing their domain sequences, Arabidopsis group I and the slime mold *WRKY* genes are more similar to rice group Ia genes than to rice group Ib genes (data

not shown). Hence, only group Ia genes, but not group Ib genes, appear to be directly derived from ancestral forms of the *OsWRKY* genes in plants. By referring to the domain feature (Fig. 2), the phylogenic tree (Fig. 3, left), and the intron pattern (Fig. 3, right), we deduced three major steps of *OsWRKY* gene evolution that involve both loss and gain of the WRKY domain (Fig. 8). Among *OsWRKY* genes, we find evidence of loss of the N terminus in the evolution of single-domain group II and IV genes from group Ia genes and independent loss of the C terminus in the evolution of a single-domain group II gene. Other group II genes may have evolved by this process. The domain loss scenario is also supported by studies of *WRKY* genes in Arabidopsis and tomato. *AtWRKY10* has only one WRKY domain, but it is clustered with the two-domain group I *AtWRKY* genes (Eulgem et al., 2000). However, its ortholog from tomato has two WRKY domains, suggesting that a recent likely loss of the N-terminal WRKY domain occurred during evolution (Rossberg et al., 2001). The domain duplication scenario is deduced based on the observation that the two zinc fingers in each group I gene are either C_2H_2-C_2H_2 or C_2HC-C_2HC type; no hybrid double zinc finger structures (C_2HC-C_2H_2 or C_2H_2-C_2HC) have been identified. Therefore, group Ib *OsWRKY* genes are likely to have evolved by intramolecular duplication of a group III WRKY domain that had already evolved the C_2HC-type zinc finger (Fig. 8). Further phylogenetic analyses based on WRKY domain DNA sequences are needed to confirm the deduced evolutionary history of this gene family.

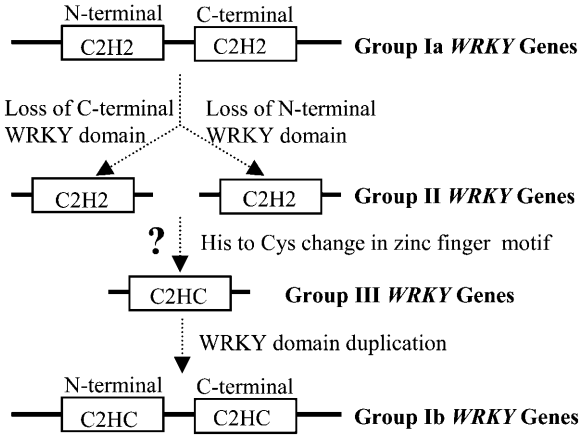

Figure 8. Deduced evolutionary framework of the *OsWRKY* gene family. An evolutionary framework of the *OsWRKY* gene family was deduced based on the phylogenetic analysis and the features of the WRKY domain. Rectangles designate the N- or C-terminal WRKY domains with a C_2H_2 or C_2HC type of zinc finger motif. Some group II *WRKY* genes are likely to be evolved from group Ia *WRKY* genes by losing either the N- or the C-terminal WRKY domain. The replacement of the conserved His residue with a Cys residue in the zinc finger motif might result in the evolution of group III *WRKY* genes. Group Ib *WRKY* genes might be evolved from group III *WRKY* genes by duplication of the C_2HC-type WRKY domain.

SA and H_2O_2 strongly induce the expression of *WRKY* genes in several plant species including rice (Kim et al., 2000; Wen et al., 2003). In Arabidopsis, 49 out of 72 tested *WRKY* genes respond to SA treatment or bacterial infection (Dong et al., 2003). The involvement of *WRKY* genes in abiotic stress response (Hara et al., 2000; Huang and Duman, 2002; Pnueli et al., 2002; Rizhsky et al., 2002, 2004; Seki et al., 2002) suggests that some of them might be responsive to ABA. Indeed, 4 Arabidopsis *WRKY* genes, *AtWRKY25* (At2g30250), *-28* (At4g18170), *-40* (At1g80840), and *-75* (At5g13080), are ABA inducible in vegetative tissues (Seki et al., 2002). In this study, 4 *OsWRKY* genes, *OsWRKY24*, *-51*, *-71*, and *-72*, were shown to be ABA responsive (Fig. 6). Multiple sequence alignment data (Z. Xie and Q.J. Shen, unpublished data) indicate these four *OsWRKY* genes are unlikely to be the orthologs of those ABA-inducible Arabidopsis *WRKY* genes. Hence, it is likely that other *WRKY* genes, which have not yet been analyzed experimentally, are also responsive to ABA. By having different temporal and spatial expression patterns, these *WRKY* genes might modulate ABA signaling in different tissues or different developmental stages.

Currently, the mechanism by which OsWRKY72 and *-77* synergistically interact with ABA to enhance the expression of *HVA22-GUS* (Fig. 7) remains unknown. Most of the published WRKY proteins bind to the cognate cis-acting element (C/T)TGAC(T/C) (Ishiguro and Nakamura, 1994; Rushton et al., 1995; Eulgem et al., 2000; Zhang et al., 2004). However, other binding sequences have been reported. For example, the barley SUSIBA2 protein binds to 3 other sequences, named SURE-a, SURE-b, and SURE-c, in addition to the W-box of the isoamylase1 promoter (Sun et al., 2003). The putative W-box has been found in the promoter regions of *HVA22* (Shen et al., 1993) and *ABF* (Choi et al., 2000). However, the 49-bp promoter fragment in the reporter construct of this study does not include this W-box. Similarly, this promoter does not contain the *Sph*I element that is bound by the C-terminal B3 domain of VP1 (Suzuki et al., 1997). Instead, only two elements are present in this promoter fragment: the ABA response element (ABRE) that is bound by ABI5 or its related bZIP proteins (Kim and Thomas, 1998; Hobo et al., 1999; Finkelstein and Lynch, 2000; Casaretto and Ho, 2003) and CE1 that is bound by an APETALA2 domain containing transcription factor ABI4 (Finkelstein et al., 1998; Niu et al., 2002). It is suggested (Shen et al., 1996) that the interaction of ABRE and CE-binding proteins might be mediated by VP1, which also interacts with 14-3-3 protein, ring (C_3HC_3-type) finger proteins, and RNA polymerase II subunit RPB5 (Schultz et al., 1998; Hobo et al., 1999; Jones et al., 2000; Kurup et al., 2000). Hence, OsWRKY72 and -77 might, as a non-DNA-binding component, form a complex with VP1, 14-3-3 protein, ring (C_3HC_3-type) finger proteins, ABI4, and ABI5 to modulate ABA signaling in aleurone cells.

In contrast, OsWRKY24 and -45 repressed ABA induction of *HVA22-GUS* (Fig. 7). If the model proposed for the OsWRKY positive regulators described above is correct, one possible repression mechanism would be the competition with the WRKY positive regulators in ABA signaling. In this regard, it is interesting to note that two rice bZIP proteins (OsZIP-2a and OsZIP-2b), which do not bind to the ABRE element by themselves, heterodimerize via the Leu zipper with EmBP-1 (another bZIP protein) and prevent it from binding ABRE. As a result, the EmBP-1 positive regulator is sequestered by these two negative regulators (Nantel and Quatrano, 1996). A similar mechanism has been found with helix-loop-helix (HLH) proteins, for which there are about 150 members in Arabidopsis (Toledo-Ortiz et al., 2003). Eighty-nine of them are predicted (Toledo-Ortiz et al., 2003) to bind to the G-box (CACGTG), which is similar to ABRE. Indeed, one of the Arabidopsis HLH proteins, *AtMYC2*, has been shown to be a transcriptional activator of ABA signaling (Abe et al., 2003). HLH proteins can function as negative regulators of basic HLH proteins by forming non-DNA-binding heterodimers with otherwise DNA-binding basic HLH proteins (Littewood and Evan, 1998). If this mode of action also holds true for the OsWRKY regulators, a different domain must be involved in dimerization because neither the positive nor the negative regulators contain a Leu zipper domain (Fig. 5; Supplemental Fig. 1) or a HLH domain. Of course, there are other possible mechanisms of repression for these repressors, as we suggested for OsWRKY71 in regulating GA signaling (Zhang et al., 2004).

In summary, the annotation of the *OsWRKY* genes has facilitated the identification of their gene structures and phylogenic relationships. In addition to SA (Chen and Chen, 2000; Dellagi et al., 2000; Du and Chen, 2000; Eulgem et al., 2000; Kim et al., 2000; Asai et al., 2002), jasmonic acid (Li et al., 2004), and GA (Rushton et al., 1992; Zhang et al., 2004), *WRKY* genes also modulate the responses of plant cells to ABA (this work). Recently, studies of parsley WRKY proteins using the chromatin immunoprecipitation technique indicate that the *PcPR-1* promoter is constitutively repressed by some WRKY proteins. Elicitor treatments induce the expression of other *WRKY* family members and the displacement of the WRKY repressors by WRKY activators (Turck et al., 2004). The WRKY repressors and activators presented in the work will help test the model and contribute to the dissection of the transcriptional complex involved in ABA signaling.

MATERIALS AND METHODS

Analyses of the Rice Genomic Sequences

About 287 Mb of rice (Oryza sativa) L. subsp. *japonica* BAC or P1 artificial chromosome sequences were downloaded from the International Rice Genome Sequencing Project (IRGSP; http://rgp.dna.affrc.go.jp/IRGSP/), which covers about 70% of the rice genome. About 361 Mb of rice *indica* contig sequences were downloaded from Beijing Genomics Institute (BGI; http://btn.genomics.org.cn/rice), which covers about 90% of the rice genome. A stand-alone version of GENSCAN (Burge and Karlin, 1997) was used for genome annotations. We used a model file for maize (*Zea mays*), which is also a monocot species that is close to rice on the taxonomy lineage because a model file trained on rice genes is not publicly available for GENSCAN. Rice peptide and coding region databases were constructed based on GENSCAN output files. Rice FL-cDNA sequences were obtained from the NCBI (http://www.ncbi.nlm.nih.gov). The sequence of rice chromosome 1 was downloaded from the Rice Genome Research Group (RGP; http://rgp.dna.affrc.go.jp/) to determine the locations of *OsWRKY* genes on chromosome 1.

HMMER software package version 2.1.1 (Sonnhammer et al., 1998) was downloaded from http://hmmer.wustl.edu/ and applied to identify *WRKY* genes in the rice peptide database. A BLAST software package (Altschul et al., 1997) was applied to compare genomic and cDNA sequences and to determine the chromosomal locations of the *OsWRKY* genes. To identify redundant sequences of the predicted *WRKY* genes, ClustalW (Thompson et al., 1994) was used to build a maximum local alignment of the predicted *WRKY* gene sequences. Then the program fRedun, which was developed in our lab using the PERL language, was applied to calculate the percentage of identity between a pair of sequences. Those with 95% or greater identity are considered redundant sequences.

Phylogenetic Analyses and Display of Gene Structures

The maximum likelihood tree was reconstructed with the WRKY domain peptide sequences using version 3.0 of the MrBayes program (Huelsenbeck et al., 2000) under the JTT model of amino acid substitutions with autocorrelated rates across sites. The Bayesian analysis was run for 200,000 generations, using 4 Markov chains with default heating values. A tree was sampled every 50 generations. The Markov chains converged rapidly, and the first 1,000 sampled trees were discarded as burn in. The majority rule consensus unrooted tree was computed based on the rest of the 3,000 sampled trees in the program PAUP*4.0 (Swofford, 2000). The statistical support values were shown above every node.

Intron phases were divided into three groups based on the position of introns within the reading frame of a gene as described in previous studies (Sharp, 1981): phase 0 introns interrupt between two codons; phase 1 introns lie between the first and second bases in a codon, and phase 2 introns interrupt a codon between the second and third bases. Gene structures were drawn to scale according to the size of exons and introns. First, exon coordinates were computed based on the revised version of *OsWRKY* genes presented in this study, and then intron phases were calculated based on exon coordinates. HMMsearch was performed by using the revised OsWRKY peptide sequences, and domain coordinates were retrieved from the HMMsearch output file. The exon and domain coordinates and intron phase information were used to graphically display the structure of *OsWRKY* genes.

Plant Materials and Incubation Conditions

Rice seeds were kindly provided by Dr. Kent McKenzie at the California Rice Experiment Station (cv M202) and Mr. Jack de Wit from DeWit Firm (cv M104; Davis, CA). Barley (*Hordeum vulgare*) cv Himalaya seeds (1998 harvest) were purchased from Washington State University (Pullman, WA). The preparation and imbibition of the embryoless half-seeds were done as described previously (Shen et al., 1993). After being imbibed for 2 d, about one-tenth of rice embryoless half-seeds were harvested as a control, and the rest of them were further incubated in petri dishes containing 20 mM $CaCl_2$, 20 mM sodium succinate without hormone (control) or with 1 μM GA_3, 20 μM ABA, or 1 μM GA_3 plus 20 μM ABA at 26°C for 4, 24, and 48 h. After harvest, rice embryoless half-seeds were immediately frozen in liquid nitrogen and stored at −80°C until further analysis.

RNA Gel-Blot Analyses

Total RNA was isolated from rice embryoless half-seeds using the Concert Plant RNA Reagent (Invitrogen, Carlsbad, CA). For northern analyses, 10 μg of total RNA were loaded into each lane, electrophoresed, and transferred onto nylon membranes. Gene-specific probes for *OsWRKY24*, *-45*, *-51*, *-71*, *-72*, *-77*, and *OsActin1* were labeled in the presence of α-^{32}P-dCTP by using the

Prime-a-Gene Labeling System (Promega, Madison, WI). Gene-specific probes for *RAmy1A* and *RAB21* were labeled with α-^{32}P-dCTP by PCR. The sequences of primers used for northern analyses were listed in Supplemental Table VI. The blots were hybridized and washed according to the method of Church and Gilbert (1984). The storage phosphor screens were scanned with a Typhoon 9410 phosphoimager (Amersham Biosciences, Piscataway, NJ).

Construct Preparations

Genomic DNA, isolated from 10-d-old rice seedlings as described previously (Zhang et al., 2004), was used for preparation of effector constructs. The genomic CRSs of the tested *OsWRKY* genes were amplified by PCR and cloned into the expression vector (Zhang et al., 2004) containing the constitutive maize ubiquitin promoter. The detailed information about primers and restriction enzyme sites used to clone the effector constructs is listed in Supplemental Table VII.

Particle Bombardment and Transient Expression Assays

Three types of DNA constructs were used in the transient expression experiments: reporter, effector, and internal control. Plasmid *HVA22-GUS* (Shen et al., 1996) was used as the reporter construct to study the function of effectors in the ABA signal transduction pathway. Plasmid *UBI-LUC*, which contains the luciferase reporter gene driven by the maize ubiquitin promoter (Bruce et al., 1989), was used as an internal control construct to normalize GUS activities of the reporter construct. The procedure of transient expression experiments with the barley aleurone system by particle bombardment has been described previously (Lanahan et al., 1992). All experiments consisted of four replicates and the entire experiment was repeated at least twice with similar results.

Sequence data from this article (OsWRKY78–OsWRKY81) have been deposited with the EMBL/GenBank data libraries under accession numbers BK005212 to BK005215.

ACKNOWLEDGMENTS

We thank Mr. Everett Cook and Ms. Tracy Hall for their excellent technical assistance. The manuscript was critically reviewed by Drs. Andrew Andres, Frank van Breukelen, and Deborah Hoshizaki in our department, and by Dr. Imre E. Somssich at Max-Plank-Institute for Plant Breeding Research, Köln, Germany. We appreciate their help. The authors are also grateful to Dr. Tuan-Hua David Ho at Washington University in St. Louis for stimulating discussions and for his generosity in providing the reporter and internal control constructs used in this study. We also thank Dr. Shoshi Kikuchi at the National Institute of Agrobiological Sciences in Japan for sharing the full-length *OsWRKY* cDNA clones prior to publication.

Received October 4, 2004; returned for revision November 11, 2004; accepted November 15, 2004.

LITERATURE CITED

Abe H, Urao T, Ito T, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2003) Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. Plant Cell **15:** 63–78

Acharya A, Ruvinov SB, Gal J, Moll JR, Vinson C (2002) A heterodimerizing leucine zipper coiled coil system for examining the specificity of a position interactions: amino acids I, V, L, N, A, and K. Biochemistry **41:** 14122–14131

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25:** 3389–3402

Asai T, Tena G, Plotnikova J, Willmann MR, Chiu WL, Gomez-Gomez L, Boller T, Ausubel FM, Sheen J (2002) MAP kinase signaling cascade in *Arabidopsis* innate immunity. Nature **415:** 977–983

Bruce WB, Christensen AH, Klein T, Fromm M, Quail PH (1989) Photoregulation of a phytochrome gene promoter from oat transferred into rice by particle bombardment. Proc Natl Acad Sci USA **86:** 9692–9696

Burge C, Karlin S (1997) Prediction of complete gene structures in human genomic DNA. J Mol Biol **268:** 78–94

Casaretto J, Ho TH (2003) The transcription factors HvABI5 and HvVP1 are required for the abscisic acid induction of gene expression in barley aleurone cells. Plant Cell **15:** 271–284

Chen C, Chen Z (2000) Isolation and characterization of two pathogen- and salicylic acid-induced genes encoding WRKY DNA-binding proteins from tobacco. Plant Mol Biol **42:** 387–396

Chen C, Chen Z (2002) Potentiation of developmentally regulated plant defense response by AtWRKY18, a pathogen-induced Arabidopsis transcription factor. Plant Physiol **129:** 706–716

Choi H, Hong J, Ha J, Kang J, Kim SY (2000) ABFs, a family of ABA-responsive element binding factors. J Biol Chem **275:** 1723–1730

Church GM, Gilbert W (1984) Genomic sequencing. Proc Natl Acad Sci USA **81:** 1991–1995

de Pater S, Greco V, Pham K, Memelink J, Kijne J (1996) Characterization of a zinc-dependent transcriptional activator from Arabidopsis. Nucleic Acids Res **24:** 4624–4631

Dellagi A, Helibronn J, Avrova AO, Montesano M, Palva ET, Stewart HE, Toth IK, Cooke DE, Lyon GD, Birch PR (2000) A potato gene encoding a WRKY-like transcription factor is induced in interactions with *Erwinia carotovora* subsp. *atroseptica* and *Phytophthora infestans* and is coregulated with class I endochitinase expression. Mol Plant Microbe Interact **13:** 1092–1101

Dong J, Chen C, Chen Z (2003) Expression profiles of the Arabidopsis WRKY gene superfamily during plant defense response. Plant Mol Biol **51:** 21–37

Du L, Chen Z (2000) Identification of genes encoding receptor-like protein kinases as possible targets of pathogen- and salicylic acid-induced WRKY DNA-binding proteins in *Arabidopsis*. Plant J **24:** 837–847

Eulgem T, Rushton PJ, Robatzek S, Somssich IE (2000) The WRKY superfamily of plant transcription factors. Trends Plant Sci **5:** 199–206

Eulgem T, Rushton PJ, Schmelzer E, Hahlbrock K, Somssich IE (1999) Early nuclear events in plant defence signalling: rapid gene activation by WRKY transcription factors. EMBO J **18:** 4689–4699

Finkelstein RR, Lynch TJ (2000) The Arabidopsis abscisic acid response gene ABI5 encodes a basic leucine zipper transcription factor. Plant Cell **12:** 599–609

Finkelstein RR, Wang ML, Lynch TJ, Rao S, Goodman HM (1998) The Arabidopsis abscisic acid response locus ABI4 encodes an APETALA 2 domain protein. Plant Cell **10:** 1043–1054

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, et al (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science **296:** 92–100

Haas BJ, Volfovsky N, Town CD, Troukhan M, Alexandrov N, Feldmann KA, Flavell RB, White O, Salzberg SL (2002) Full-length messenger RNA sequences greatly improve genome annotation. Genome Biol **3:** RESEARCH0029

Hara K, Yagi M, Kusano T, Sano H (2000) Rapid systemic accumulation of transcripts encoding a tobacco WRKY transcription factor upon wounding. Mol Gen Genet **263:** 30–37

Hinderhofer K, Zentgraf U (2001) Identification of a transcription factor specifically expressed at the onset of leaf senescence. Planta **213:** 469–473

Hobo T, Kowyama Y, Hattori T (1999) A bZIP factor, TRAB1, interacts with VP1 and mediates abscisic acid-induced transcription. Proc Natl Acad Sci USA **96:** 15348–15353

Huang J, Zhao Q, Mooney SM, Lee FS (2002) Sequence determinants in hypoxia-inducible factor-1alpha for hydroxylation by the prolyl hydroxylases PHD1, PHD2, and PHD3. J Biol Chem **277:** 39792–39800

Huang T, Duman JG (2002) Cloning and characterization of a thermal hysteresis (antifreeze) protein with DNA-binding activity from winter bittersweet nightshade, *Solanum dulcamara*. Plant Mol Biol **48:** 339–350

Huelsenbeck JP, Rannala B, Larget B (2000) A Bayesian framework for the analysis of cospeciation. Evolution Int J Org Evolution **54:** 352–364

Ishiguro S, Nakamura K (1994) Characterization of a cDNA encoding a novel DNA-binding protein, SPF1, that recognizes SP8 sequences in the 5′ upstream regions of genes coding for sporamin and beta-amylase from sweet potato. Mol Gen Genet **244:** 563–571

Johnson CS, Kolevski B, Smyth DR (2002) *TRANSPARENT TESTA GLABRA2*, a trichome and seed coat development gene of Arabidopsis, encodes a WRKY transcription factor. Plant Cell **14:** 1359–1375

Jones HD, Kurup S, Peters NC, Holdsworth MJ (2000) Identification

and analysis of proteins that interact with the Avena fatua homologue of the maize transcription factor VIVIPAROUS 1. Plant J **21:** 133–142

Karrer EE, Litts JC, Rodriguez RL (1991) Differential expression of alpha-amylase genes in germinating rice and barley seeds. Plant Mol Biol **16:** 797–805

Kikuchi S, Satoh K, Nagata T, Kawagashira N, Doi K, Kishimoto N, Yazaki J, Ishikawa M, Yamada H, Ooka H, et al (2003) Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice. Science **301:** 376–379

Kim CY, Lee SH, Park HC, Bae CG, Cheong YH, Choi YJ, Han C, Lee SY, Lim CO, Cho MJ (2000) Identification of rice blast fungal elicitor-responsive genes by differential display analysis. Mol Plant Microbe Interact **13:** 470–474

Kim SY, Thomas TL (1998) A family of novel basic leucine zipper proteins binds to seed-specification elements in the carrot Dc3 gene promoter. J Plant Physiol **152:** 607–613

Kurup S, Jones HD, Holdsworth MJ (2000) Interactions of the developmental regulator ABI3 with proteins identified from developing Arabidopsis seeds. Plant J **21:** 143–155

Lagace M, Matton DP (2004) Characterization of a WRKY transcription factor expressed in late torpedo-stage embryos of Solanum chacoense. Planta **219:** 185–189

Lanahan MB, Ho THD, Rogers SW, Rogers JC (1992) A gibberellin response complex in cereal alpha-amylase gene promoters. Plant Cell **4:** 203–211

Li J, Brader G, Palva ET (2004) The WRKY70 transcription factor: a node of convergence for jasmonate-mediated and salicylate-mediated signals in plant defense. Plant Cell **16:** 319–331

Littewood T, Evan GI (1998) Helix-Loop-Helix Transcription Factors, Ed 3. Oxford University Press, New York

McElroy D, Rothenberg M, Reece KS, Wu R (1990) Characterization of the rice (Oryza sativa) actin gene family. Plant Mol Biol **15:** 257–268

Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW (2003) Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. Plant Cell **15:** 809–834

Mundy J, Chua N-H (1988) Abscisic acid and water-stress induce the expression of a novel rice gene. EMBO J **7:** 2279–2286

Nantel A, Quatrano RS (1996) Characterization of three rice basic/leucine zipper factors, including two inhibitors of EmBP-1 DNA binding activity. J Biol Chem **271:** 31296–31305

Niu X, Helentjaris T, Bate NJ (2002) Maize ABI4 binds coupling element1 in abscisic acid and sugar response genes. Plant Cell **14:** 2565–2575

Pnueli L, Hallak-Herr E, Rozenberg M, Cohen M, Goloubinoff P, Kaplan A, Mittler R (2002) Molecular and biochemical mechanisms associated with dormancy and drought tolerance in the desert legume *Retama raetam*. Plant J **31:** 319–330

Pu WT, Struhl K (1991) The leucine zipper symmetrically positions the adjacent basic regions for specific DNA binding. Proc Natl Acad Sci USA **88:** 6901–6905

Rizhsky L, Davletova S, Liang H, Mittler R (2004) The zinc finger protein Zat12 is required for cytosolic ascorbate peroxidase 1 expression during oxidative stress in Arabidopsis. J Biol Chem **279:** 11736–11743

Rizhsky L, Liang H, Mittler R (2002) The combined effect of drought stress and heat shock on gene expression in tobacco. Plant Physiol **130:** 1143–1151

Robatzek S, Somssich IE (2001) A new member of the *Arabidopsis* WRKY transcription factor family, *At*WRKY6, is associated with both senescence- and defence-related processes. Plant J **28:** 123–133

Robatzek S, Somssich IE (2002) Targets of AtWRKY6 regulation during plant senescence and pathogen defense. Genes Dev **16:** 1139–1149

Rossberg M, Theres K, Acarkan A, Herrero R, Schmitt T, Schumacher K, Schmitz G, Schmidt R (2001) Comparative sequence analysis reveals extensive microcolinearity in the lateral suppressor regions of the tomato, Arabidopsis, and Capsella genomes. Plant Cell **13:** 979–988

Rushton PJ, Hooley R, Lazarus CM (1992) Aleurone nuclear proteins bind to similar elements in the promoter regions of two gibberellin-regulated alpha-amylase genes. Plant Mol Biol **19:** 891–901

Rushton PJ, Macdonald H, Huttly AK, Lazarus CM, Hooley R (1995) Members of a new family of DNA-binding proteins bind to a conserved *cis*-element in the promoters of alpha-*Amy2* genes. Plant Mol Biol **29:** 691–702

Savkur RS, Burris TP (2004) The coactivator LXXLL nuclear receptor recognition motif. J Pept Res **63:** 207–212

Schultz TF, Medina J, Hill A, Quatrano RS (1998) 14-3-3 proteins are part of an abscisic acid-VIVIPAROUS1 (VP1) response complex in the Em promoter and interact with VP1 and EmBP1. Plant Cell **10:** 837–847

Seeler JS, Dejean A (2003) Nuclear and unclear functions of SUMO. Nat Rev Mol Cell Biol **4:** 690–699

Seki M, Narusaka M, Ishida J, Nanjo T, Fujita M, Oono Y, Kamiya A, Nakajima M, Enju A, Sakurai T, et al (2002) Monitoring the expression profiles of 7,000 *Arabidopsis* genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. Plant J **31:** 279–292

Sharp PA (1981) Speculations on RNA splicing. Cell **23:** 643–646

Shen Q, Ho THD (1995) Functional dissection of an abscisic acid (ABA)-inducible gene reveals two independent ABA-responsive complexes each containing a G-box and a novel cis-acting element. Plant Cell **7:** 295–307

Shen Q, Uknes SJ, Ho THD (1993) Hormone response complex of a novel abscisic acid and cycloheximide inducible barley gene. J Biol Chem **268:** 23652–23660

Shen Q, Zhang P, Ho THD (1996) Modular nature of abscisic acid (ABA) response complexes: composite promoter units that are necessary and sufficient for ABA induction of gene expression in barley. Plant Cell **8:** 1107–1119

Sonnhammer EL, Eddy SR, Birney E, Bateman A, Durbin R (1998) Pfam: multiple sequence alignments and HMM-profiles of protein domains. Nucleic Acids Res **26:** 320–322

Sun C, Palmqvist S, Olsson H, Boren M, Ahlandsberg S, Jansson C (2003) A novel WRKY transcription factor, SUSIBA2, participates in sugar signaling in barley by binding to the sugar-responsive elements of the iso1 promoter. Plant Cell **15:** 2076–2092

Suzuki M, Kao CY, McCarty DR (1997) The conserved B3 domain of VIVIPAROUS1 has a cooperative DNA binding activity. Plant Cell **9:** 799–807

Swofford D (2000) PAUP*: Phylogenetic Analysis Using Parsimony. Sinauer Associates, Sunderland, MA

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res **22:** 4673–4680

Tiwari SB, Hagen G, Guilfoyle TJ (2004) Aux/IAA proteins contain a potent transcriptional repression domain. Plant Cell **16:** 533–543

Toledo-Ortiz G, Huq E, Quail PH (2003) The Arabidopsis basic/helix-loop-helix transcription factor family. Plant Cell **15:** 1749–1770

Turck F, Zhou A, Somssich IE (2004) Stimulus-dependent, promoter-specific binding of transcription factor WRKY1 to its native promoter and the defense-related gene PcPR1-1 in parsley. Plant Cell **16:** 2573–2585

Ulker B, Somssich IE (2004) WRKY transcription factors: from DNA binding towards biological function. Curr Opin Plant Biol **7:** 491–498

Verger A, Perdomo J, Crossley M (2003) Modification with SUMO. A role in transcriptional regulation. EMBO Rep **4:** 137–142

Wen N, Chu Z, Wang S (2003) Three types of defense-responsive genes are involved in resistance to bacterial blight and fungal blast diseases in rice. Mol Genet Genomics **269:** 331–339

Xu YH, Wang JW, Wang S, Wang JY, Chen XY (2004) Characterization of GaWRKY1, a cotton transcription factor that regulates the sesquiterpene synthase gene (+)-delta-cadinene synthase-A. Plant Physiol **135:** 507–515

Yu D, Chen C, Chen Z (2001) Evidence for an important role of WRKY DNA binding proteins in the regulation of *NPR1* gene expression. Plant Cell **13:** 1527–1540

Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, et al (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science **296:** 79–92

Zhang ZL, Xie Z, Zou X, Casaretto J, Ho TH, Shen QJ (2004) A rice WRKY gene encodes a transcriptional repressor of the gibberellin signaling pathway in aleurone cells. Plant Physiol **134:** 1500–1513

Zhou W, Ryan JJ, Zhou H (2004) Global analyses of sumoylated proteins in Saccharomyces cerevisiae: induction of protein sumoylation by cellular stresses. J Biol Chem **279:** 32262–32268